ALGEBRA

LECTURES DELIVERED TO POST-GRADUATE STUDENTS OF CALCUTTA UNIVERSITY

BY

FRIEDRICH WILHELM LEVI, DR. PHIL. NAT.
HARDINGE PROFESSOR

PART I

SYSTEMS OF LINEAR EQUATIONS

PUBLISHED BY THE
UNIVERSITY OF CALCUTTA
1936

PRINTED AND PUBLISHED BY BHUPENDRALAL BANERJEE
AT THE CALCUTTA UNIVERSITY PRESS, SENATE HOUSE, CALCUTTA.

Reg. No. [illegible]—August, 19[illegible]—[illegible]

PREFACE

On introducing a new course of lectures in Algebra I realized after delivering a few lectures that the students of this country should have in their hands a book covering the whole subject-matter of the lectures. In order to make my lectures successful I had no other alternative than to write a text-book and to publish it in different parts as quickly as possible.

So a provisory edition of this text-book is taken in hand. References to the original papers, examples, explanation of details, everything that causes delay of publication had to be omitted in these "lectures." Later on a full text-book on Algebra will be published.

In placing this fascicule in the hands of the students, I offer my heartiest thanks to our energetic Vice-Chancellor, SYAMAPRASAD MOOKERJEE, Esq., M.A., B.L., Barrister-at-Law, M.L.C., without whose sympathetic co-operation this publication would not have come into being. I thank also the Calcutta University Press for having printed this paper in a very short time under difficult circumstances.

Proofs and manuscripts have been revised by Mr. R. C. Bose, M.A., Mr. S. K. Bhar, M.Sc., and especially by Mr. A. C. Choudhury, M.Sc. If the reader do not find many offences against the spirit of the English language, he should be thankful to these three young colleagues of mine.

CALCUTTA, ASUTOSH BUILDING, F. W. LEVI.
August, 1936.

§ 1. Introduction.

The problem we have to deal with in this first part is the following:

Given $m(n+1)$ arbitrary numbers

$$\begin{matrix} a_0, a_1, a_2, \ldots\ldots\ldots, a_n \\ b_0, b_1, b_2, \ldots\ldots\ldots, b_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ k_0, k_1, k_2, \ldots\ldots\ldots, k_n \end{matrix}$$

called the "*coefficients*," we are required to find out the numbers

$$x_1, x_2, \ldots\ldots\ldots x_n \qquad \ldots \qquad \ldots \quad (1)$$

satisfying the m conditions:

$$\left.\begin{matrix} a_1x_1 + a_2x_2 + \ldots\ldots\ldots + a_nx_n = a_0 \\ b_1x_1 + b_2x_2 + \ldots\ldots\ldots + b_nx_n = b_0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ k_1x_1 + k_2x_2 + \ldots\ldots\ldots + k_nx_n = k_0 \end{matrix}\right\} \quad \ldots \qquad \ldots \quad (2)$$

The system (2) is called a *system of linear equations*. A linear equation is called *homogeneous*, if its term on the right side in (2) equals 0, and also the system (2) is called homogeneous, if every equation is homogeneous, i.e., if $a_0 = b_0 \ldots = k_0 = 0$. Every system of numbers (1) satisfying (2) is called a solution of (2).

At first we will try to get the solutions in the most simple cases.

I. $n = m = 1$ $\qquad a_1x = a_0$

(α) $a_1 \neq 0$. $\qquad$ Solution $x = a_0 : a_1$

(β) $a_1 = 0$ $\quad a_0 \neq 0$. $\qquad$ No solution.

(γ) $a_1 = a_0 = 0$. $\qquad$ Solution x is arbitrary.

Remark: Division is permissible, if and only if the denominator is not equal to zero. If the denominator is not a constant, but a function, it is necessary to treat separately the cases in which this function vanishes. In the lecture, examples of this kind will be given.

II. $n=2, m=1$ $\quad a_1x_1+a_2x_2=a_0$

(α) $a_1\neq 0, a_2\neq 0$. Solutions (x_1, x_2); one of the numbers is arbitrary, the other is defined by it.

(β) $a_1\neq 0, a_2=0$. Solutions $(a_0:a_1, x_2)$ x_2 is arbitrary.

($\beta\beta$) $a_1=0, a_2\neq 0$. „ $(x_1, a_0:a_2)$ x_1 „

(γ) $a_1=a_2=0, a_0\neq 0$. No solution.

(δ) $a_1=a_2=a_0=0$. Solutions x_1 and x_2 are arbitrary.

III. $n=1, m=2$

$$a_1x=a_0$$
$$b_1x=b_0$$

(α) $a_1b_0-b_1a_0\neq 0$. No solution.

(β) $a_1b_0-b_1a_0=0$, $(a_1, b_1)\neq(0, 0)$.

Solutions: $a_0:a_1$, or $b_0:b_1$, or $a_0:a_1=b_0:b_1$

if $a_1\neq 0$ $\quad b_1\neq 0$ $\quad a_1\neq 0$, $b_1\neq 0$.

(γ) $a_1=b_1=0$, $(a_0, b_0)\neq(0, 0)$. No solution.

(δ) $a_1=b_1=a_0=b_0=0$. Solution x is arbitrary.

IV. $n=2, m=2$

$$a_1x_1+a_2x_2=a_0$$
$$b_1x_1+b_2x_2=b_0$$

$$a_1b_2-a_2b_1=\Delta$$

$$a_0b_2-a_2b_0=\Delta_1$$

$$-a_0b_1+a_1b_0=\Delta_2$$

Necessary conditions for solutions (x_1, x_2) are $\Delta x_1 = \Delta_1$
$\Delta x_2 = \Delta_2$

(α) $\Delta \neq 0$. One solution $(\Delta_1 : \Delta, \Delta_2 : \Delta)$

(β) $\Delta = 0$, $(\Delta_1, \Delta_2) \neq (0, 0)$. No solution.

(γ) $\Delta = \Delta_1 = \Delta_2 = 0$, $a_1 b_2 (= b_1 a_2) \neq 0$.

Every solution of the first equation is also a solution of the second.—See II (α).

(δ) $a_1 = a_2 = 0$, $a_0 \neq 0$. No solution.

(δ') $b_1 = b_2 = 0$, $b_0 \neq 0$. No solution.

(ϵ) $a_1 = a_2 = a_0 = 0$. Every solution of the second equation is a solution of the system.—See II.

(ϵ') $b_1 = b_2 = b_0 = 0$. Every solution of the first equation is a solution of the system.—See II.

(ζ) $a_1 = b_1 = \Delta_1 = 0$. Solutions.—See III, β, γ, δ.

(ζ') $a_2 = b_2 = \Delta_2 = 0$. Solutions.—See III, β, γ, δ.

§ 2. The Homogeneous System belonging to an Arbitrary System of Linear Equations.

From the examples given in § 1 we may realize that even in the most simple cases a great number of different sub-cases has to be considered. The number of these sub-cases seems to increase infinitely with n and m. For avoiding these difficulties we will follow a way very characteristic of mathematical thoughts. We will suppose that solutions have been found out, and we will enquire about the connection between those solutions. By considering the properties of the system of all solutions we will get a general theory of the systems of linear equations including also different ways for finding out the solutions of a given system.

If $(x_1, x_2, \ldots, x_n) = (w_1, w_2, \ldots, w_n)$ and
$(x_1, x_2, \ldots, x_n) = (u_1, u_2, \ldots, u_n)$ are solutions of (2), then
$(x_1, x_2, \ldots, x_n) = (w_1 - u_1, \ldots, w_n - u_n)$ is a solution of

$$\begin{array}{c} a_1 x_1 + \ldots\ldots + a_n x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ k_1 x_1 + \ldots\ldots + k_n x_n = 0 \end{array} \qquad (2/H)$$

and conversely, if $(x_1, \ldots, x_n)$ is a solution of (2), and $(y_1, \ldots, y_n)$ a solution of (2/H), then $(x_1 + y_1, \ldots, x_n + y_n)$ is a solution of (2). Therefore the following theorem holds:

Theorem 1. Starting from an arbitrary solution of (2) we will get all solutions by addition of the solutions of (2/H).

The homogeneous system (2/H) belongs to the system (2). For solving (2) we have to find out all solutions of (2/H) and an arbitrary one of (2). For that purpose the introduction of a new notion is convenient.

§ 5. The *n-vectors*

Definition 1: An ordered set of n numbers is called an n-vector.

$$\alpha = (a_1, \ldots\ldots, a_n) \tag{3}$$

The n numbers a_i defining the n-vector are called its co-ordinates. As this set is an ordered one, the vector will generally be changed by the interchange of the co-ordinates.

Examples: (1) The co-efficients of an arbitrary equation of (2/H) define an n-vector; it is called the "vector of that equation," and also the "vector of that row."

(2) The co-efficients of an arbitrary column of (2/H) define an m-vector, the "vector of that column."

(3) The solution (1) of (2) defines an n-vector, the "vector of the solution."

(4) Vectors in the plane (the space), in the sense this word is ordinarily used, are 2-vectors (3-vectors).

(5) Let n be the number of the customers of a bank; the balances of the customers are the co-ordinates of an n-vector representing the actual state of the bank. The reader may interpret the vector-addition defined below for this example.

Definition 2: The product of a number c and the vector α is an n-vector.

$$c\alpha = (ca_1, \ldots ca_n). \tag{4}$$

Definition 3: The sum of α and $\beta = (b_1, \ldots, b_n)$ is

$$\alpha + \beta = (a_1 + b_1, \ldots, a_n + b_n). \tag{5}$$

From these definitions it follows:

$$\begin{aligned} \alpha+\beta &= \beta+\alpha && \text{commutative law,} \\ \alpha+(\beta+\gamma) &= (\alpha+\beta)+\gamma && \text{associative law,} \\ c\,(\alpha+\beta) &= c\alpha+c\beta && 1^{st} \text{ distributive law,} \\ (c_1+c_2)\alpha &= c_1\alpha+c_2\alpha && 2^{d} \text{ distributive law.} \end{aligned} \tag{6}$$

As these laws hold, we can use the notations of sum of n-vectors in the same manner as it is to be used for numbers:

$$\sum c_i\alpha^i = (\sum c_i a_1^i, \ldots, \sum c_i a_n^i) \tag{7}$$

$i = 1, \ldots, m$; c being arbitrary numbers; $\alpha^i = (a_1^i, \ldots, a_n^i)$

being arbitrary n-vectors.

The vector $-1\cdot\alpha$ is called the *negative* of α and written $-\alpha$. The addition of $-\alpha$ is the inverse operation to the addition of α. As in elementary arithmetics this inverse operation is called *subtraction* and written by the sign $-$. Therefore:

$$\beta+(-\alpha) = \beta-\alpha\,. \tag{8}$$

The following special n-vectors will often be used:

$$\begin{aligned} 0 &= (0, \ldots, 0) && \text{Zero-vector} \\ e^1 &= (1, 0, \ldots, 0) && 1^{st} \text{ Unit-vector} \\ e^2 &= (0, 1, 0, \ldots, 0) && 2^{d} \text{ Unit-vector} \\ &\cdots\cdots\cdots\cdots \\ e^n &= (0, 0, \ldots, 0, 1) && n^{th} \text{ Unit-vector.} \end{aligned} \tag{9}$$

Formulae:

$$\begin{aligned} \alpha-\alpha &= 0 \\ c\cdot 0 &= 0 \\ \alpha &= \sum a_i e^i \end{aligned} \tag{10}$$

§ 4. Vector-spaces.

Definition 3: The n-vector $\Sigma c_i \alpha^i$ is *dependent* on the n-vectors α^i.

Definition 4: The n-vectors $\alpha^1, \ldots\ldots, \alpha^m$ are said to be *independent* if none of them is dependent on the $m-1$ others and $m>1$. A single n-vector is independent, if it is $\neq 0$.

Definition 5: The set of all vectors dependent on the α^i is called the *vector-space* generated by the α^i.

Definition 6: A set of independent n-vectors generating a vector-space is called its *Basis*.

Theorems concerning vector-spaces:

1. The n-vectors α^i are independent if and only if for every system of numbers $(c_1, \ldots, c_m) \neq (0, \ldots, 0)$ we have $\Sigma c_i \alpha_i \neq 0$.

Proof. 1. If $c_i \neq 0$, α^i is dependent on the other α^i. 2. If α^j is dependent on the other α^i, then $\alpha^j = \Sigma d_\mu \alpha^\mu$, and therefore $\Sigma c_i \alpha^i = 0$, for $c_j = -1$, $c_\mu = d_\mu$. The theorem is evident in the case of a single n-vector.

2. If $\beta^1, \ldots, \beta^r$ belong to a vector-space V, every n-vector κ dependent on the β^i belongs also to V.

Proof. Let V be generated by $\alpha^1 \ldots \alpha^m$, then from $\kappa = \Sigma k_i \beta^i$, $\beta^i = \Sigma b^i_j \alpha^j$ follows $\kappa = \Sigma c_j \alpha^j$, where $c_j = \Sigma k_i b^i_j$.

3. Every vector-space containing n-vectors $\neq 0$ has a basis.

Proof. If the vectors α^i generating V are not independent, α^j may depend on the other $m-1$ generating n-vectors. From 2 it follows that these vectors generate also V. On repeating—if it is necessary—this reduction we will get after a finite number of steps a subset of the α^j generating V composed of independent vectors, *i.e.*, a basis of V.

4. If $\alpha^1, \ldots, \alpha^m$ is a basis of V, $\beta = \Sigma c_i \alpha^i$, and $c_j \neq 0$, then we will get a new basis of V on replacing α^j by β.

Proof. The vector-space generated by β and the $\alpha^{i \neq j}$ is contained in V. On the other hand it contains $\beta - \sum_{k \neq j} c_k \alpha^k = c_j \alpha^j$ and therefore α^j. We have to show that the m n-vectors β and $\alpha^{i \neq j}$ are independent.

Let $d\beta + \sum_{k \neq i} d_k \alpha^k = 0$, on replacing β by its value $\sum c_k \alpha^k$ we get a vanishing linear function of the α^k whose coefficients vanish, as the α^k are independent. The coefficient of α^i is dc_i; as $c_i \neq 0$, it follows: $d=0$. As the α^k are independent, the d_k are vanishing. Therefore the α_k and β are independent; so they form a basis of V.

5. If $\alpha^1, \ldots, \alpha^m$ is a basis of V, and $\beta^1, \ldots, \beta^t$ are independent vectors in V, then we get a new basis on replacing t suitable α^i by the β, and $t \leq m$ holds.

Proof. From 4 it follows that a suitable α can be replaced by β^1. Let $\alpha^1, \ldots, \alpha^{m-r}, \beta^1, \ldots, \beta^r$ be a basis of V, $r<t$ then we can express β^{r+1} by these m vectors, and in this expression the coefficient of at least one α do not vanish; therefore this α can be replaced by β^{r+1}. Therefore we can continue replacing an α by a β till $r=t$, or $r=m$. In the last case all β's must depend on the $\beta^1, \ldots, \beta^m$, and therefore $t=m$ in this case. Generally $t \leq m$.

Definition 7: The maximum number of independent vectors of V is called the Rank of V.

6. The number of the vectors of an arbitrary basis of V equals the rank of V. (Therefore every basis of V has the same number of vectors.)

Proof. From 5 it follows that there cannot be more independent n-vectors in V than an arbitrary basis has elements.

7. If every n-vector of V belongs to V', but not every n-vector of V' belongs to V, then is the rank of V less than the rank of V'.

Proof. Let $\alpha^1, \ldots, \alpha^r$ be a basis of V, and β a vector of V' not contained in V, then is r the rank of V, but the rank of V' is at least $r+1$ because V' contains $r+1$ independent elements α^i, β.

8. The rank of a vector-space of n-vectors is at most n.

Proof. The n unit-vectors (9) generate a vector-space V' in which the co-ordinates are arbitrary numbers; therefore V' contains every n-vector, and 8 follows from 7.

9. Between $p > n$ of n-vectors there exists always a linear equation with non-vanishing coefficients.

Proof. If these n vectors are independent they will generate an n-vector-space of rank $p > n$.

1[illegible] A system A of n vectors with the property that the sum of two arbitrary elements of A, and also the product of an arbitrary element of A with an arbitrary real number belongs to A, is a vector space.

Proof. From [illegible] it follows that in A there exists a maximum number r [illegible] of independent vectors. The vector space generated by such r independent vectors is identical with A. If $r = 0$, A contains only the n-vector 0.

4. The Vector Spaces Connected with a System of Homogeneous Linear Equations

Theorem II. The solutions of 2/II form a vector space X.

Proof. From $\sum a_i x_i = \sum b_i x_i = \dots = \sum k_i x_i = 0$

and $\sum a_i y_i = \sum b_i y_i = \dots = \sum k_i y_i = 0$ it follows

$$\sum a_i c x_i = \sum b_i c x_i = \dots = \sum k_i c x_i = 0 \quad \text{and}$$

$$\sum a_i (x_i + y_i) = \sum b_i (x_i + y_i) = \dots = \sum k_i (x_i + y_i) = 0$$

Therefore the solutions satisfy the conditions of a vector space given in §4, 10.

Theorem III. To every vector space X there exists a vector space V such that every n-vector of X is a solution of a linear homogeneous equation if and only if the vector of this equation is a vector of V.

Proof. The vectors of X are solutions of the equation defined by the n-vector [illegible]. If they are solutions of the equations defined by [illegible] and by [illegible] then they are also solutions of the equations defined by [illegible] and by [illegible]. From §4, 10 it follows that the set of the vectors [illegible] form a vector space.

The Theorems II and III show that our problem (2/II) is closely connected with two vector spaces X and V. The vectors of 2/II generate a vector space V' and every vector of V' is also a vector of V. If V and V' were not identical, then V should be of higher rank. In that case two vector spaces of equation vectors of different rank would define the same

vector space V. In other words there would exist any vector λ independent of the vectors of the equations (2/H), such that every solution of (2/H) is also a solution of the equation generated by λ. We will see later that this case is impossible.

§ 6. The Basis of a Vector-Space. The Method of Steinitz

For further consideration of linear equations we need know some more properties of vector spaces.

11. If $\alpha, \beta, \ldots, \kappa$ are n vectors generating a vector space V, $a \neq 0$ and b are arbitrary numbers, then $a\alpha + b\beta, \beta, \ldots, \kappa$ generate also V.

Proof. As V includes $a\alpha + b\beta$ and the vector-space generated by $a\alpha + b\beta, \beta, \ldots, \kappa$ includes α, the two vector spaces are identical.

12. If the n vectors $\alpha, \beta, \ldots, \kappa$ generating V are not all equal to 0, then we will get another system of n vectors generating V by omitting the n-vectors equal to 0.

Proof. The n-vector 0 is dependent on the other ones.

13. The n-vectors

$$
\begin{aligned}
\beta^1 &= (1, 0, \ldots, 0, b^1_{r+1}, \ldots, b^1_n) \\
\beta^2 &= (0, 1, \ldots, 0, b^2_{r+1}, \ldots, b^2_n) \\
&\ldots\ldots\ldots \\
\beta^r &= (0, 0, \ldots, 1, b^r_{r+1}, \ldots, b^r_n)
\end{aligned}
\tag{11}
$$

are independent.

Proof. Let $\sum c_i \beta^i = 0$; the coordinates of the vector on the left are equal to the coordinates on the right hand. Therefore $c_1 = c_2 = \ldots = c_r = 0$.

14. V may be an arbitrary vector space including n-vectors $\neq 0$; starting from an arbitrary finite system of n-vectors generating V and using the methods of 11 and 12 we will get after a finite number of suitable steps a basis of V differing from the n-vectors of formula 11 at most by a permutation of the co-ordinates of all vectors.

Remark. The methods 11 and 12 can be considered as operations practised on the scheme of co-ordinates written on the right hand in formula

11. It is therefore useful to [illegible] 14 as a property of that scheme. For this purpose we introduce a notation often used in mathematics.

Definition 5. The system of the co-ordinates of m ordered n-vectors is called a Matrix M. The co-ordinates of the i-th vector form the i-th row of M, [illegible] co-ordinates of the vectors form the [illegible] column of M [illegible] ($1 \leq i \leq m$, $1 \leq k \leq n$). M = 0 means that every co-ordinate of M vanishes.

The operations given by 11 and 12 may be shortly expressed by the words row addition and [illegible] of rows. Using these words we get a theorem (14') equivalent to (14).

14'. By the operations of row addition and [illegible] of the rows every matrix M ≠ 0 can be transformed to a matrix of the kind (14) or to a matrix differing from it only by a permutation of the columns. The proof of 14' and therefore also the proof of 14 [illegible] be made by different steps. To simplify the description of these transformations it may be understood that every row that would appear should be omitted automatically without mentioning that operation. Every row contains therefore at least one co-ordinate ≠ 0. [illegible]

the rows of the matrix (1), (2) [illegible]

the co-ordinates [illegible] [illegible] (12)

the columns [illegible] $\langle 1\rangle, \langle 2\rangle, \ldots \langle n\rangle$

These symbols [illegible] constant values [illegible] change at every step of transformation [illegible] denotes the first row at every step, also $\langle k\rangle$ the k-th column, [illegible] being an arbitrary number, etc.

Observing these [illegible] the different steps of the transformation may be described as follows:

[illegible] may [illegible] the smallest number such that [illegible] ≠ 0

by the row addition [illegible] → [illegible] (1)

[illegible] becomes −1 ;

[illegible] by the row addition [illegible] → [illegible] ≠ [illegible]

[illegible] becomes 0

By these procedures the column $<i_1>$ has been swept out, i.e. one of the numbers has been made -1, the others vanish. We continue on sweeping out

2. a) i_2 may be the smallest number such that $[i_2, 2] \neq 0$, then $i_2 \neq i_1$, by row addition $(i_2) \longrightarrow$ [illegible] (i_2)

$[i_2, 2]$ becomes -1 and $[i_1, 2]$ is not changed.

2 b) By the row addition $(k) \longrightarrow (k) + [k, 2](i_2)$

for every $k \neq i_2$, $[k, 2]$ becomes 0 and $<i_1>$ is not changed.

Now the columns $<1>$ and $<2>$ are swept out. We continue this sequence of operations. After $2(q-1)$ steps the columns

$<1>, <2>, \ldots, <q-1>$ may be swept out, i.e.

$[i_s, s] = -1$, $[k, s] = 0$ for $s = 1, \ldots, q-1$, $k \neq i_s$, $i_1, \ldots, i_{q-1}$

being different numbers. If after these steps (and the eventual omitting of 0 rows) the matrix has more than $q-1$ rows:

q. a) i_q may be the smallest number for which $[i_q, q] \neq 0$,

by $(i_q) \longrightarrow (-1/[i_q, q])(i_q)$, $[i_q, q]$ becomes -1.

q. b) By $(k) \longrightarrow (k) + [k, q](i_q)$, $k \neq i_q$, $[k, q]$ becomes 0.

The rows $<i_1>, \ldots, <i_{q-1}>$ are not changed by these transformations.

By mathematical induction it follows that the method can be continued till the number of columns swept out equals the number of the remaining rows. After this transformation

$<1>, \ldots, <r>$ may be swept out, i.e. $[i_s, s] = -1$, $[k, s] = 0$

for $s = 1, \ldots, r$, $k \neq i_s$. By a permutation of the columns transforming $i_s \longrightarrow s$ the matrix will take the form (1).

Corollary. Starting from an arbitrary system of n-vectors generating $V \neq 0$ it is always possible to get a basis of V.

Proof. Let m be the number of the generating vectors. It is always possible to sweep out the matrix by at most m^2 row additions and omissions of at most $m-1$ rows. As the n-vectors of the remaining rows are independent, they form a basis.

§ 7. Solution of Systems of Homogeneous Linear Equations by the Method of "Sweep-out"

Theorem IV. Let V be the vector space generated by the n vectors of [illegible] H), X the vector space of the solutions of [illegible] H), and V' the vector space [illegible] (13)

then V' = V, rank (V) [illegible] rank [illegible]

Theorem V. It is possible to find out a basis of X by a limited number of row additions and omissions of 0 rows.

Proof of the Theorems IV and V. Every [illegible] of (2 H) is a solution of the equations belonging to the n vectors of V and vice versa. To [illegible] these [illegible] out the solutions of the equations belonging to a basis of V. As it was proved by (4) it is also possible to find out such a basis by the method of sweep-out. This basis can be taken in the form (1) only by a permutation $i_k \longrightarrow k$ of the columns. If we write therefore $x_{i_k} = y_k$, $k = 1, \ldots, n$

then our problem will be reduced to the following

$$y_1 + b^1_{r+1} y_{r+1} + \cdots + b^1_n y_n = 0$$

$$\cdots\cdots\cdots\cdots$$

$$y_r + b^r_{r+1} y_{r+1} + \cdots + b^r_n y_n = 0 \tag{14}$$

$$y_i = \sum_{j=1}^{n-r} b^i_j \, y_{r+j} \quad [illegible]$$

The $y_{r+1}, \ldots, y_n$ can take arbitrary independent values and $y_1, \ldots, y_r$ are uniquely determined by them. There exists therefore a (1, 1) correspondence between the vector space of all $(n-r)$ vectors $(y_{r+1}, \ldots, y_n)$, and the vector space of all solutions $y_1, \ldots, y_n$;

$$\begin{matrix} (1, 0, \ldots, 0) & & (b^1_{r+1}, \ldots, b^r_{r+1}, 1, 0, \ldots, 0, 0) \\ \cdots\cdots & \text{and} & \cdots\cdots \\ (0, 0, \ldots, 0, 1) & & (b^1_n, \ldots, b^r_n, 0, 0, \ldots, 0, 1) \end{matrix}$$

are corresponding vectors. As this correspondence is invariant for vector addition and multiplication with a number, and the $(n-r)$ vectors on the right form a basis of the solutions (14), we will get a basis of X by the permutation $k \longrightarrow i_k$ of the coordinates.

Therefore Theorem V is true, and

$$\text{Rank}\,(V) + \text{rank}\,(X) = n$$

holds. As X is also the vector space of all solutions of (1) we can replace V' by V in the formula, hence it holds. Every vector of V' belongs to V and these vector spaces have the same rank, then for (§4, 2) V and V' are identical.

§ 8. Solutions of Non-homogeneous Linear Systems

In order to solve (2) it would be useful to consider [illegible] and

$$\begin{aligned} a_1 x_1 + \cdots + a_n x_n + a x_0 &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ k_1 x_1 + \cdots + k_n x_n + k x_0 &= 0 \end{aligned} \qquad (15)$$

the generating $(n+1)$ vectors being called

$a = (a_1, \dots, a_n, a), \dots\dots, k = (k_1, \dots, k_n, k)$.

Theorem VI. The system (2) is solvable if and only if the vector space V generated by the n vectors $a, \dots, k$ and the vector space V' generated by the $n+1$ vectors $a, \dots, k$ have the same rank. If (2) is not solvable, then rank V' = 1 + rank V.

Proof. As every linear homogeneous relation between the $a, \dots, k$ holds when $a, \dots, k$ are replaced by 0, the rank r of V is not greater than the rank r' of V'. Let X and X' be the vector spaces generated by the solutions of (1) and (15), then rank $X = n - r$, rank $X' = n + 1 - r'$. If $(x_1, \dots, x_n)$ is a solution of (2), then $(x_1, \dots, x_n, 1)$ is a solution of (15). Therefore rank $X' \geq$ rank $X + 1 \geq r' \geq r$. If $r' = r + 1$, rank X' = rank X, then (15) has only solutions with $x_0 = 0$, and (2) has no solutions. If $r' = r$, rank X' > rank X and there exists a solution $(x_1, \dots, x_n, x_0)$, $x_0 \neq 0$ of (15). Hence $\left(\frac{x_1}{x_0}, \dots\dots, \frac{x_n}{x_0}\right)$ is a solution of (2).

Theorem VII. If there are solutions of (2), we will get them by "sweep out".

Proof. We can get a basis of the solutions X' of (15) by the method of sweep out. If there is a solution in which $x_0 \neq 0$, it must also be a basis $(n+1)$ vector of this kind. From this $(n+1)$ vector we will get a solution of (2) on dividing the coordinates by x_0.

As 0 is a solution of (2) if and only if the equations are homogeneous, the solutions form a vector space only in this case. In general the set of all solutions will be described by the following theorem.

Theorem VIII. If ξ and η are solutions of (2) and $s+t=1$ then $s\xi+t\eta$ is also a solution. Conversely if a set W of n-vectors has the property that with two n-vectors ξ and η the n-vector $s\xi+t\eta$ belongs to W, then W is the set of the solutions of a system (2).

Proof. From $\sum a_{ik} x_k = a_{i0}$, $\sum a_{ik} y_k = a_{i0}$ it follows that $\sum a_{ik}(s x_k + t y_k) = (s+t) a_{i0} = a_{i0}$. As the same holds for the other equations of (2) the first proposition is proved. — If χ, λ, ν are n-vectors of W then [illegible] and [illegible] belongs to W. Hence [illegible]. If ν is a fixed element of W, the set of all differences $(\chi-\nu)$, $(\lambda-\nu)$ is a vector space X [illegible]. From Theorem III it follows that X is the set of the solutions of a system of linear homogeneous equations. This system may be (2 H). If $\nu = (p_1, \ldots, p_n)$ and we define $a_{10}, \ldots$ by $a_{i0} = \sum p_k a_{ik}$, [illegible] then by Theorem I W is the set of all solutions of (2).

THE METHOD OF ORTHOGONALISATION

By the previous theorems the problem proposed in the introduction has been solved completely. It will always be the principal part of the calculation to find out a basis of the vector space X. We may do it by the method of sweep-out but it is important to have other suitable methods for it. In this section the method of orthogonalisation will be treated; its advantage is that we will get at the same time a basis of X and a basis of V, both in a special form.

In the previous, and in the following paragraphs there is no restriction about the numbers which should be used.[1] However in this section we will suppose that all numbers used are real.

Definition 9: The *Scalar product* $S.\alpha\beta$ of two n-vectors $\alpha=(a_1, \ldots, a_n)$ and $\beta=(b_1, \ldots, b_n)$ is:

$$S.\alpha\beta = S.\beta\alpha = \sum a_i b_i \quad \ldots \quad (16)$$

Definition 10: The *Length* of α is a number $|\alpha| \geq 0$, defined by

$$|\alpha|^2 = S.\alpha\alpha \quad (17)$$

[1] Choosing the notation introduced in Chapter II. The theorems and the proofs hold for an arbitrary field of characteristic [illegible]

Definition [illegible] α and on γ if [illegible]

$$S.\alpha(\beta+\gamma)=S.\alpha\beta+S.\alpha\gamma$$
$$S.c\alpha\beta \quad =cS.\alpha\beta \qquad (18)$$

$$[illegible] \; |S.\alpha\beta| \leq [illegible] \leq |\alpha|\cdot|\beta| \quad \text{(Cauchy's inequality)},$$

$$|\alpha+\beta| \leq |\alpha|+|\beta|$$

Definition 11. If $S.\alpha\beta=0$, α and β are orthogonal.

Remark. In the lectures [illegible] the significance of these notions and formulae will be explained.

The coordinates of α can be expressed as scalar products

$$a_i = S.\alpha\alpha^i \qquad \ldots \quad (19)$$

Definition 12. The vectors $\beta^1, \ldots, \beta^n$ form an *Orthogonal System* if they satisfy the conditions:

$$S.\beta^i\beta^i=1$$
$$S.\beta^i\beta^j=0 \qquad i \neq j. \qquad \ldots \quad (20)$$

Properties of orthogonal systems

1. If the β^i form an orthogonal system, they are independent.

Proof. From $0=\Sigma c_i\beta^i$ it follows

$$0=\beta^k\Sigma c_i\beta^i=c_k \text{ for } k=1, \ldots, n.$$

2. If the β^i form an orthogonal system, and γ is independent of the β^i, then there is an n-vector β^{n+1} such that γ is dependent on $\beta^1, \ldots, \beta^{n+1}$, and these vectors form an orthogonal system.

Proof. $\lambda=\gamma-\Sigma S.\beta^i\gamma\,\beta^i$, and $\beta^{n+1}=|\lambda|^{-1}\lambda$, then there is for $k=1, \ldots, n$, $0=S.\beta^k\lambda=S.\beta^k\beta^{n+1}$, $S.\beta^{n+1}\beta^{n+1}=1$.

3. If A is a vector space of rank m containing a vector space A' of rank $r<m$, then there exists a basis of A forming an orthogonal system $\beta^1, \ldots, \beta^r, \ldots, \beta^m$, such that $\beta^1, \ldots, \beta^r$ form a basis of A'.

Proof. Let β^1 be an arbitrary n-vector of A' of the length 1. If $r>1$ there is in A' an n-vector independent of β^1, and therefore there is an orthogonal system β^1, β^2. By repeated use of this construction we get $\beta^1, \ldots, \beta^r$ in A' and by continuing this method in A we will get $\beta^{r+1}, \ldots, \beta^m$

Theorem IX. Let V be the vector space generated by the vectors of the equations (2.1) and X the vector space of the solutions of (2.1). Then there exists an orthogonal system $a^1, \ldots, a^r, \xi^1, \ldots, \xi^{n-r}$, such that $a^1, \ldots, a^r$ form a basis of V, and $\xi^1, \ldots, \xi^{n-r}$ form a basis of X.

Proof. The connection between the vector spaces V and X is that every n-vector of X is orthogonal to every n-vector of V and vice versa. Hence if we construct an orthogonal system of n n-vectors, $a^1, \ldots, a^r, \xi^1, \ldots, \xi^{n-r}$, such that the first r n-vectors form a base of V, the $\xi^1, \ldots, \xi^{n-r}$ are orthogonal to the n-vectors of V and form therefore solutions of (2.1). An arbitrary n-vector can be expressed by $\lambda = \sum \alpha_i a^i + \sum \beta_j \xi^j$. λ is orthogonal to the n-vectors of V if and only if for $k = 1, \ldots, r$ $0 = a^k \lambda = \alpha_k$ holds. $\xi^1, \ldots, \xi^{n-r}$ is therefore a basis of X.

Remark. The proof of Theorem IX does not apply to the Theorems IV–VIII and the method of sweeping out. Theorem IV follows directly from Theorem IX.

§ 10. Substitution and Elimination

The method of substitution. In order to get the solutions of (2) we can also reduce the problem to $m-1$ equations and $n-1$ unknowns. Let x_{i_1} be the first unknown for which the coefficient of the 1st equation does not vanish, then:

$$x_{i_1} = -\sum_{k \neq i_1} \frac{a_{1k}}{a_{1i_1}} x_k + \frac{a_1}{a_{1i_1}}$$

On substituting this value in the other $m-1$ equations we get $m-1$ equations with $n-1$ unknowns. If some of these equations become identical they will be omitted. The rows of the new equations are

$$(l_i - [i_1, k] : a_i(1),$$

hence the substitution of x_{i_1} is nothing more than the sweeping out of the column $\langle i_1 \rangle$. On continuing this procedure the column $\langle i_2 \rangle$ will be swept out in the rows $3, \ldots, m$, also $\langle i_2 \rangle$ in the rows (3), etc., the 0 rows, as they belong to identical equations being always omitted. The result will be that the columns $\langle i_1 \rangle, \ldots, \langle i_r \rangle$ will be swept out in the rows below $[i, \cdot]$. The last of these equations gives us the possibility to express x_{i_r} by the remaining $n-r$ unknowns. The substitution of this

value in the [illegible] is nothing but the sweep-out of the column [illegible] in the rows above the row [illegible]. On repeating this procedure by the substitution of [illegible] expressed by the [illegible] unknown we sweep out the corresponding column in the upper rows. Hence the method of substitution is not quite different from the method of sweep-out.

[illegible]. The equations (2.1) will be solved when we have [illegible] vectors belonging to [illegible] are independent on the vectors belonging to (2.1), we should multiply the equations (2.1) with [illegible] numbers and add so that the sum of these equations becomes [illegible] we eliminate the unknown x_q [illegible] x_1 to get the first row of [illegible] to get the second row, etc. By this procedure we intend to find out at once what is done step by step by using the method of sweep-out. The rank of [illegible] and also the factors we need for the elimination are dependent only on the given coefficients. We have [illegible] that dependence. For this purpose we need the determinants.

II. PRINCIPAL PROPERTIES OF THE DETERMINANTS

Definition [illegible]. A function

$$L = L(a^1, \ldots, a^n) \qquad (21)$$

of n [illegible] vectors

$$a^1, \ldots, a^n$$

is called a *Determinant* [illegible]

$$[illegible] = \det a, \qquad (22)$$

if the following conditions are satisfied:

(a) $L(a^1, \ldots, a^n) = [illegible]$ for [illegible] $= 1, \ldots, n$

(b) L will not change by replacing

$$a_i \longrightarrow \lambda a_k + a_i, \text{ for } i \neq k, \qquad \ldots (23)$$

(c) $L(e^1, \ldots, e^n) = 1$.

It will be proved later on that a function of this kind exists, and that it is uniquely defined by the conditions (a), (b), (c). Now we will consider

the properties of a function [illegible] satisfying the above conditions

1. If $a^i = 0$, $L = 0$.

Proof. $0 = 0 \cdot L = L(a^1, \ldots, 0 \cdot a^i, \ldots) = L(a^1, \ldots)$ [illegible]

2. L will not be changed by replacing [illegible] $\rightarrow$ [illegible] if [illegible] and c is an arbitrary number.

Proof. $L =$ [illegible]

$=$ [illegible]

$=$ [illegible]

3. On interchanging [illegible] L is replaced by $-L$.

Proof. $L = L(a^1,$ [illegible] $)$

$= L($ [illegible] $)$

$=$ [illegible]

$=$ [illegible]

4. If [illegible] is dependent on the [illegible] other vectors, $L = 0$.

Proof. [illegible] can be replaced by 0. Hence $L = 0$.

5. $L(a^{i_1}, \ldots, a^{i_n}) = 1$, if $i_1, \ldots, i_n$ is an even permutation

$= -1$, if $i_1, \ldots, i_n$ is an odd permutation.

$= 0$, if the indices are not different.

These formulae follow directly from [illegible] 3 and 4.

6. $L(a^1, \ldots, a^i + \sum b_\nu a^\nu, \ldots, a^n) = L(a^1, \ldots) + L$ [illegible]

Proof. Either a^i is dependent on the other n vectors a^ν or these $n-1$ vectors are not independent or the n vectors are all independent. In the first case $L = 0$ on the left side [illegible] may be omitted and therefore the equation holds. In the 2d case each of the three functions equals zero. In the 3d case [illegible] is dependent on the [illegible] $= \sum b_\nu$ [illegible]. By applying 2 on both sides of the equation the left side becomes $(1 + b_i)L$, and the right side becomes $L + b_i L$, therefore 6 holds.

7. If [illegible] and D, is the determinant we get by [illegible], then [illegible]

Proof. To prove [illegible] we have to use [illegible]

8. Let L_i^j be the determinant we get by [illegible] and let $a^i = \sum_j a_j^i e^j$, then

$$L = \sum a_j^i L_i^j \qquad \ldots \qquad \ldots \qquad (24)$$

The theorem follows [illegible] from [illegible]

$$9. \; L = \sum \pm a_1^{\cdot} \ldots a_n^{\cdot} \qquad \ldots \qquad \ldots \qquad \ldots \qquad \ldots \qquad (25)$$

the summation has to be made [illegible] the permutations [illegible] being valid for even and [illegible] for odd permutations.

Proof. On using [illegible] we get [illegible] hence from 5 follows the theorem.

Theorem X. The determinant [illegible] exists and is uniquely defined by every ordered set of n vectors [illegible]

Proof. As it was stated in 9 the value of the determinant—if there exists one—cannot be different from [illegible]. Therefore we have to prove that [illegible] has the properties [illegible] and [illegible] are [illegible] satisfied we will prove [illegible]. The right side of [illegible] is a function of [illegible] may be called $D(a^1, \ldots, a^n)$. By interchanging two of the vectors of the argument the even permutations become odd and [illegible] therefore D is changed to $-D$. Hence if [illegible] then $D = -D = 0$. From [illegible] it follows that [illegible] and [illegible] the [illegible] (b) is satisfied and the theorem holds.

§ 12. [illegible] Properties of [illegible] Determinants

The columns of [illegible] form [illegible]

$$e_1 = [illegible]$$

$$e_n = [illegible]$$

[illegible] (2.1)

[illegible] determinant [illegible] not [illegible] changed [illegible] are [illegible] changed by the [illegible].

Proof. Let [illegible] be the permutation inverse to [illegible] then these permutations are [illegible] both even or both odd [illegible]

[illegible] and

[illegible]

[illegible] the [illegible] if we replace the a^i by the a_i.

These theorems are [illegible] of (2.1).

11. If we replace [illegible] by [illegible] the determinant becomes independent of the co-ordinates of a_q.

Proof. On replacing [illegible] by [illegible] the product [illegible] becomes [illegible] hence [illegible] vanishing [illegible] the determinant becomes independent of the co-ordinates [illegible]

[illegible] determinant we get on replacing [illegible] by [illegible] being different and [illegible] being different will be called

(2.9)

12. Let [illegible] be the determinant we get on replacing [illegible] the arguments [illegible] by [illegible] then

$$[illegible]$$

Proof. From [illegible] we see that [illegible] is not changed if we replace the [illegible] by [illegible]. From 10 and 11 it follows also that [illegible] is not changed if we replace the [illegible]. In both cases we get a determinant with

[illegible]

the [illegible]

$$[\ldots] = \ldots \text{ for } [\text{illegible}]$$

$$= 1 \quad \text{for } (i, k) = (p_\nu, q_\nu) \quad \nu = 1, \ldots, m \qquad (29)$$

$$= 0 \quad \text{for } i = p_\nu, \; k \neq q_\nu$$

$$\text{or } i \neq p_\nu, \; k = q_\nu.$$

[illegible] By a permutation of the [illegible] and a permutation of the [illegible] when these permutations are either both even or both odd [illegible] changed to its negative value, when the two permutations are of different kind.

Proof. To every permutation [illegible] belongs a corresponding permutation of the [illegible]. Hence the theorem holds.

If $n > 2$ [illegible] every [illegible] an even permutation.

[illegible]

[illegible] from the given [illegible] every [illegible] combinations the following theorem holds.

[illegible]

= 1 when [illegible] all [illegible]

[illegible] permutation of [illegible]

[illegible] are not [illegible] different numbers.

Proof. Let $p_1, \ldots$ [illegible] an even permutation of $1, \ldots, m$. The factors of each term of [illegible] may be developed as functions of the [illegible] according to [illegible]. The first factor is composed of [illegible] minors of degree [illegible] the second factor by [illegible] minors of degree [illegible]. Hence [illegible] is the sum of [illegible] of the

form [illegible] (32)

$a_1, \ldots, a_m$ is an arbitrary permutation [illegible], and $r_1, \ldots, r_{n-m}$ an arbitrary permutation of [illegible]. The sign + has to be taken if and only if both permutations are of the same kind [illegible], if

$$a_1, \ldots, a_m, r_1, \ldots, r_{n-m} \qquad (33)$$

is an even permutation of $1, \ldots, n$,

and therefore also an even permutation of [illegible] $r_1, \ldots, r_{n-m}$.

Every permutation of [illegible], the number of the monomials in [illegible] and therefore the sum becomes L. In the case of an odd permutation the sign of every monomial [illegible] is changed and therefore we get -1. When the numbers j and r are not all different [illegible] a determinant with two identical rows and from [illegible] it follows that it is vanishing.

The relation between the two factors of each term of [illegible] is a reciprocal one; they are called cofactors. We can extend the notion of cofactor and the formula [illegible] to the case [illegible] by taking

$$L_1^1 = a_1^1, \; L_2^1 = a_1^2, \; L_1^2 = a_2^1, \; L_2^2 = a_2^2$$

Often it is used to be [illegible] $m = n - 1$ [illegible]

$$\sum a_i^j L_k^i = 0, \text{ for } i \neq j$$

$$= L, \text{ for } i = j. \qquad (34)$$

From [illegible] it follows that [illegible] holds also when the upper and the lower indices are interchanged; hence we get

$$\sum a_i^j L_j^k = 0, \text{ for } i \neq j$$

$$= L, \text{ for } i = j. \qquad (34')$$

If an arbitrary Matrix D has [illegible] [illegible] $= a_k^i$ we will write

$$D = [illegible]$$

On [illegible] some rows and some columns of D, we get new matrices

$$D_{i_1 \ldots}^{k_1 \ldots} = [illegible], \quad a_k^i = b_k^i$$

If m [illegible] the [illegible] define [illegible], called the [illegible] of D of order m

$$\det D^{h_1, \ldots, h_m}_{k_1, \ldots, k_m} = \det [illegible] . \qquad ([illegible])$$

15. If $L^{h_1, \ldots, h_m}_{k_1, \ldots, k_m}$ and $L^{h_1, \ldots}_{l_1, \ldots}$ [illegible] are cofactors, then

$$L^{h_1, \ldots, h_m}_{k_1, \ldots} = \det D^{h_1, \ldots}_{k_1, \ldots} [illegible]$$

Proof. The replacement [illegible] by which we get L [illegible] from [illegible] [illegible] every [illegible] having the factor [illegible] and in the remaining [illegible] the factor will be replaced by [illegible]. The [illegible] of L [illegible] are therefore just the [illegible] of a determinant composed of the rows [illegible] and the columns $< t_1 >$, [illegible] $< t_{n-m} >$. For getting the proper sign we must realise that the permutation of the $n - m$ rows is of the same character (even or odd) as the permutation of the $n - m$ columns if and only if the permutation of the original [illegible] rows is of the same character as the permutation of the original n columns. This condition is however satisfied by the definition of the cofactors.

16. If all minors of D of order m vanish, then the minors of higher order vanish also.

Proof. Using the formula [illegible] we can develop an arbitrary minor of order $m + l$ as a homogeneous linear function of degree m with [illegible] [illegible] being minors of order l.

Definition 14. If there is a non-vanishing minor of D of order r but every minor of order $r + 1$ vanishes, r is called the *rank* of D.

17. The matrix D, the vector space generated by its rows and the vector space generated by its columns have all the same rank.

Proof. If r is the rank of the vector space of the row vectors, there is a linear homogeneous relation between every set of $r + 1$ row vectors, and this relation holds also if we cut away some of the columns. Hence from [illegible] it follows, that the minors of order $r + 1$ are all vanishing, and

the rank of D is at most r. However, if R is the rank of D every minor of order > R is vanishing and a minor of order R is not vanishing. Without restriction of the generality we suppose $D^{1 \dots R}_{1 \dots R} \neq 0$. Let (s) be an arbitrary row.

In $D^{1 \dots R\; s}_{1 \dots R\; m+1}$ the cofactors of [illegible] $_{m+1}$, …, [illegible] $_{m+1}$ will be called respectively

$$A_1, \dots, A_R, \; A_s = D^{1 \dots R}_{1 \dots R} \neq 0.$$

Hence the coordinates of $A_1 a^1 + \dots + A_R a^R + A_s a^s$ are either determinants with two equal rows or minors of D of order R + 1 and therefore vanishing. Hence [illegible] every [illegible] depends on $a^1, \dots, a^R$, hence the rank of the vector space is at most R. As we have seen in the first part of the proof the rank of the vector space is not smaller than R and therefore both ranks are equal. The rank of D is not changed, if the rows and the columns of D are interchanged; hence the rank of D equals also the rank of the vector space generated by the columns of D.

§ 13. [illegible]

Theorem XI. The equations (2) are solvable if and only if the matrices

$$M = \begin{pmatrix} a_1 & \dots & a_n \\ \vdots & & \vdots \\ k_1 & \dots & k_n \end{pmatrix} \quad \text{and} \quad \bar{M} = \begin{pmatrix} a_1 & \dots & a_n & a_0 \\ \vdots & & & \vdots \\ k_1 & \dots & k_n & k_0 \end{pmatrix} \qquad (30)$$

have the same rank.

Proof. From (28) it follows that the ranks of the matrices are the same as the ranks of the vector spaces V and $\bar{V}$ of Theorem VI. Hence the Theorem XI follows from the Theorem VI.

In order to get the rank r of a matrix by the help of determinants it is not necessary to calculate each minor; it is sufficient to state that one minor of order r is not vanishing and that all minors of order r + 1 vanish. The rank does not change by row addition or by column addition; it is often useful to simplify matrices by these operations.

Corollary. If the matrices M and $\bar{M}$ of (30) have the same rank then the matrix formed by an arbitrary set of rows [illegible] of $\bar{M}$ has the same rank as a matrix formed by the same rows of M.

Proof: As the non-homogeneous equations belonging to M and M are solvable, the rows (p_1) ..... (p_t) define solvable equations, and therefore the two matrices defined by these t rows have the same rank.

Theorem XII. If the matrices M and M of Theorem XI have the same rank r, and a non-vanishing minor of M has the rows (p_1) ..... (p_r), then the solutions of (2) are identical with the solutions of the equations belonging to these rows.

Proof: From the Corollary it follows that the rank of the matrix formed by the (p_1) ..... (p_r) is also r. The solutions of the homogeneous equations defined by the r rows form a vector-space X of rank $n-r$, including all solutions of (2/II). As the rank of the vector-space of the solutions of (2/II) is also $n-r$, it follows from 14, 7 that every solution of the r linear homogeneous equations is also a solution of (2/II). Hence from Theorem I it follows that we get every solution of (2) by adding all vectors of X to an arbitrary solution of the r non-homogeneous equations, *i.e.*, the solutions of these equations are just the solutions of (2).

By the Theorems XI and XII the problem of solving a system of linear equations by elimination has been reduced to the following: Find out the solutions of

$$\begin{aligned} a^1_1 x_1 + \dots + a^1_r x_r + a^1_{r+1} x_{r+1} + \dots + a^1_n x_n &= a^1_0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a^r_1 x_1 + \dots + a^r_r x_r + a^r_{r+1} x_{r+1} + \dots + a^r_n x_n &= a^r_0 \end{aligned} \tag{37}$$

when the determinant det A of the matrix

$$A = \begin{pmatrix} a^1_1 \dots\dots, a^1_r \\ \dots\dots\dots\dots \\ a^r_1 \dots\dots, a^r_r \end{pmatrix}$$ is not vanishing.

Theorem XIII. Let A^i_k be the co-factors of a^i_k in A, then we get the solutions of (37) by

$$\det A \; x_h + \sum_i a^i_{r+1} A^i_h x_{r+1} + \dots + \sum_i a^i_n A^i_h x_n = \sum_i a^i_0 A^i_h \tag{38}$$

$h = 1, \dots, r$.

Proof: By multiplying the equations (37) respectively by the co-factors A^i_h (h being constant), and adding, we get the equations (38).

These equations are therefore necessary conditions for the solutions of (37). The rows of (38)—including the right sides of the equations—are dependent on the rows of (37). The rank of the matrix composed of the rows of (37) and (38) is therefore also r. The determinant formed by the first r columns of (38) is a power of det A and therefore $\neq 0$. Hence from Theorem XII it follows that the solutions of (38) are identical with the system of all solutions of (37) and (38), and therefore (38) is also a sufficient condition for the solutions of (37).

§ 14. LINEAR TRANSFORMATIONS.

Let

$$A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \dots & \dots & \dots \\ a_1^n & \dots & a_n^n \end{pmatrix} \tag{39}$$

be a matrix with n rows and n columns.

The row-vectors are called $\alpha^1, \dots, \alpha^n$,

the column-vectors are called $\alpha_1, \dots, \alpha_n$.

We consider the equations

$$\sum_k a_k^i x_k = y_i \ , \ i = 1, \dots, n. \tag{40}$$

To every n-vector $\xi = (x_1, \dots, x_n)$ corresponds an n-vector $\eta = (y_1, \dots, y_n)$; (40) is called a *linear transformation*, and we will express it by

$$\xi \longrightarrow \eta \ .$$

Theorem XIV. By a linear transformation (40) the n-vectors ξ are transformed to the n-vectors of a vector-space, whose rank equals the rank of A.

Proof: If $\xi^1 \longrightarrow \eta^1$, $\xi^2 \longrightarrow \eta^2$, then $\xi^1 + \xi^2 \longrightarrow \eta^1 + \eta^2$

and $c\xi^1 \longrightarrow c\eta^1$, for every number c.

Hence from §4. 10 it follows, that η form a vector-space H. Two n-vectors ξ are transformed to the same n-vector η, if and only if the difference is transformed to O, *i.e.*, if the difference belongs to the vector-space Z of the solutions of the homogeneous system belonging to (40). Using the methods of §4, 7, or §8, 3, we will find out a basis $\zeta^1, \dots, \zeta^{n-r}$, $\xi^1, \dots, \xi^r$ of the vector-space of ξ, so that the $n-r$ first n-vectors form

a basis of Z. If $\xi^i \longrightarrow \eta^i$, an arbitrary n-vector is transformed: $\Sigma c_i \xi^i + \Sigma d_k \zeta^k \longrightarrow \Sigma c_i \eta^i$, and from the definition of ζ^k it follows that this n-vector vanishes if and only if $c_1 = \ldots = c_r = 0$. Therefore $\eta^1, \ldots, \eta^r$ form a basis of H, and r is the rank of H.

Theorem XV. A representation of the n-vectors ξ by the n-vectors $f(\xi)$ with the properties: $f(\xi^1 + \xi^2) = f(\xi^1) + f(\xi^2)$, $f(c\xi) = cf(\xi)$, is a linear transformation.

Proof. Let $f(e^k) = \beta_k = \Sigma b_k^i e^i$, $\xi = \Sigma x_k e^k$, then $f(\xi) = \Sigma b_k^i x_k e^i$.

Hence the representation is a linear transformation with the matrix $((b_k^i))$.

If $y_i = \Sigma a_i^k x_k$, $x_k = \Sigma b_k^j z_j$, then $y_i = \Sigma a_i^k b_k^j z_j = \Sigma d_i^j z_j$, where

$$d_i^j = \Sigma a_i^k b_k^j = \text{S}.\alpha^i \beta_j, \tag{41}$$

α^i being the row-vectors of $\text{A} = ((a_i^k))$, and β_j the column-vectors of $\text{B} = ((b_k^j))$. The matrix $\text{D} = ((d_i^j))$ is called the product

$$\text{D} = \text{A} \cdot \text{B} \tag{42}$$

If we consider 3 matrices A, B, C, and their products $(\text{A} \cdot \text{B}) \cdot \text{C}$ and $\text{A} \cdot (\text{B} \cdot \text{C})$, we get in both cases a matrix with the co-ordinates $g_i^l = \sum_{k,j} a_i^k b_k^j c_j^l$; hence:

Theorem XVI. The associative law holds for the multiplication of matrices.

§15. Decomposition of Matrices.

A matrix $\text{D} = ((d_i^k))$, $d_i^i = d_i$, $d_i^k = 0$, when $i \neq k$, is called a *Diagonal-matrix*. A matrix $\text{E}_{rs}(\lambda) = ((e_i^k))$, $e_i^i = 1$, $e_r^s = \lambda$, and every other $e_i^k = 0$, is called an *Elementary-matrix* (defined only for $r \neq s$). If A has the row-vectors $\alpha^1, \ldots, \alpha^n$, and the column-vectors $\alpha_1, \ldots, \alpha_n$, then

D·A has the row-vectors $d_1 \alpha^1, \ldots, d_n \alpha^n$

A·D has the column-vectors $d_1 \alpha_1, \ldots, d_n \alpha_n$

$\text{E}_{rs}(\lambda) \cdot \text{A}$ has the row-vectors $\beta^i = \alpha^i$,

for $i \neq r$ $\qquad \beta^r = \alpha^r + \lambda \alpha^s$ (43)

$\text{A} \cdot \text{E}_{rs}(\lambda)$ has the column-vectors $\gamma_i = \alpha_i$,

for $i \neq s$ $\qquad \gamma_s = \alpha_s + \lambda \alpha_r$.

Theorem XVII. An arbitrary matrix A of n rows and n columns is a product $A = P_1 \cdot D \cdot P_2$, where P_1 and P_2 are products of elementary-matrices, and D is a diagonal-matrix.

Proof. From (43) it follows that the theorem is identical with the following: A can be transformed to a diagonal-matrix by row-additions $a^r \longrightarrow a^r + \lambda a^s$, and by column-additions $a_r \longrightarrow a_r + \lambda a_s$. In order to prove this proposition, we use the method of "sweep-out" in a little modified manner. If every element of A vanishes, A is a diagonal-matrix; if all elements do not vanish, we can make $[1, 1] \neq 0$ by row-additions and column-additions of the type mentioned above, and by the same kind or operations we can sweep out the first row and the first column. On continuing this procedure we get a diagonal-matrix.

Theorem XVIII. The determinant of $A \cdot B$ is equal to the product of the determinants of A and B, *i.e.*,

$$\det (A \cdot B) = \det A \cdot \det B.$$

Proof. From (43) it follows, that the determinant of a matrix does not change, when the matrix is multiplied with an elementary-matrix, and therefore the determinant does not change, when the matrix is multiplied with a product of elementary-matrices. Hence, if $B = P_1 \cdot D \cdot P_2$, $\det B = \det D = d_1 \ldots d_n$. On the other hand $\det (A \cdot B) = \det (A \cdot P_1 \cdot D)$ holds. As we get $A \cdot P_1 \cdot D$ from $A \cdot P_1$ by multiplying the columns with $d_1 \ldots d_n$, we get: $\det (A \cdot B) = \det (A \cdot P_1 \cdot D) = \det (A \cdot P_1) \cdot d_1 \ldots d_n = \det A \det B$.